Staff : **KAN MIN-YEN**
Department : **COMPUTER SCIENCE**

ARTICLE IN JOURNAL

INTERNATIONAL REFEREED

PUBLISHED

1. YE*, S, T S Chua*, M Kan and Long Qiu, "Document concept lattice for text understanding and summarization". *INFORMATION PROCESSING & MANAGEMENT*, 43, no. 6, part. - (- on -) (- 2007): 1643-1662. (United States).
    Publication no. 0356824

2. CUI*, H, M Kan and T S Chua, "Soft Pattern Matching Models for Definitional Question Answering". *ACM Transactions on Information Systems*, 25, no. 2 (2007). (United States).
    Publication no. 0340125

3. Kan*, M and Y F TAN*, "Record Matching in Digital Library Metadata". *COMMUNICATIONS OF THE ACM*, 51, no. 2 (Regular) (2007): 91-94. (United States).
    Publication no. 0340120

4. Lu*, W and M Kan, "Supervised Categorization of Javascript using Program Analysis Features". *INFORMATION PROCESSING & MANAGEMENT*, 43, no. 2 (Regular) (2006): 431-444. (United States). (Leading Rank).
    Publication no. 0314876

5. Elhadad*, N, M Kan, J Klavans and K R Mckeown, "Customization in a Unified Framework for Summarizing Medical Literature". *ARTIFICIAL INTELLIGENCE IN MEDICINE*, 33, no. 2 (2005): 179-198. (Netherlands). (Rank: Leading Journal).
    Publication no. 0187765

@ - indicates journal classification is based on the proposed classification of the author(s)

BOOK REVIEW

PUBLISHED

1. Kan, M. Review of "Introduction to Digital Libraries", by G. G. Chowdhury and Sudatta Chowdhury. *J. of American Society for Information Science and Technology*, 55, no. 2 (2004): 178-179. Hoboken, NJ.
Publication no. 0187741

CONFERENCE PAPER

INTERNATIONAL

ORAL PRESENTATION

PUBLISHED

1. TAN*, Y F, E Elmacioglu, M Kan and D Lee, "Efficient Web-Based Linkage of Short to Long Forms". *International Workshop on the Web and Databases (WebDB 2008)* (2008). New York: ACM Press. (International Workshop on the Web and Databases (WebDB), 13 Jun 2008, Vancouver, Canada)
Publication no. 0377023

2. Councill, IG, CL Giles and M Kan*, "ParsCit: An open-source CRF reference string parsing package.". *Language Resources and Evaluation Conference 2008* (2008). Paris: European Language Resources Association. (Language Resources and Evaluation Conference, 26 - 30 May 2008, Marrakesh, Morocco)
Publication no. 0376565

3. Bird, S, R Dale, BL Dorr, B Gibson, M Joseph, M Kan*, D Lee, B Powley, DR Radev and Y F TAN, "The ACL Anthology Reference Corpus: A Reference Dataset for Bibliographic Research in Computational Linguistics". *Language Resources and Evaluation Conference 2008* (2008). Paris: European Language Resources Association. (Language Resources and Evaluation Conference, 26 - 30 May 2008, Marrakesh, Morocco)
Publication no. 0376332

4. QIU*, L, M Kan and T S Chua, "Modeling Context in Scenario Template Creation". *International Joint Conference on Natural Language Processing 2008* (2008). Hyderabad: Asian Federation of Natural Language Processing. (International Joint Conference on Natural Language Processing, 20 - 24 Jan 2008, Hyderabad, India)
Publication no. 0377117

5. NGUYEN*, TD and M Kan, "Keyphrase Extraction in Scientific Publications". *International Conference on Asian Digital Libraries* (2007): 317-326. Hanoi: ICADL. (International Conference on Asian Digital Libraries, 10 - 13 Dec 2007, Hanoi, Vietnam)
Publication no. 0352353

6. Elmacioglu*, E, M Kan, D Lee and Y Zhang, "Web Based Linkage". *Proceedings of the Workshop on Web Information and Data Management 2007* (2007): 121-128. New York: ACM Press. (Workshop on Web Information and Data Management, 9 Nov 2007 - 9 Nov 2009, Lisboa, Portugal)
Publication no. 0352352

7. Setiawan*, H, M Kan* and H Li, "Ordering Phrases with Function Words". *Association of Computational Linguistics (ACL 07)* (2007). New York: Association for Computational Linguistics. (Association of Computational Linguistics, 25 - 28 Jun 2007, Prague, Czech Republic)
Publication no. 0340692

8. GOZALI*, J and M Kan, "A Rich OPAC User Interface with AJAX". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340135

9. Kan, M, "SlideSeer: A digital library of aligned document and presentation pairs". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340134

10. YAN*, S, D Lee, M Kan* and CL Giles, "Adaptive Sorted Neighborhood Methods for Efficient Record Linkage". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340133

11. Nguyen*, V and M Kan*, "Functional Faceted Web Query Analysis". *WWW 2007* (2007). Banff: ACM Press. (World Wide Web, 8 - 12 May 2007, Banff, Canada)
Publication no. 0350704

12. LIN*, Z, T S Chua, M Kan*, W S Lee, L QIU and S YE, "NUS at DUC 2007: Using Evolutionary Models of Text". *DUC 2007* (2007). Rochester: Association for Computational Linguistics. (Document Understanding Conference, 26 - 27 Apr 2007, Rochester, United States)
Publication no. 0350701

13. LIN*, Z and M Kan, "Timestamped Graphs: Evolutionary Models of Text for Multi-document Summarization". *Proceedings of TextGraphs-2* (2007). Rochester: Association for Computational Linguistics. (Graph-based Methods for Natural Language (TextGraphs-2), 26 - 27 Apr 2007, Rochester, United States)
Publication no. 0340129

14. Iskandar*, D, Y Wang, M Kan and H Li, "Syllabic Level Automatic Synchronization of Music Signals and Text Lyrics". *ACM Multimedia* (2006). New York: ACM Press. (ACM Multimedia, 23 - 27 Oct 2006, Santa Barbara, California, United States) (Rank Other (Short paper)).
Publication no. 0314882

15. Neo, Shi Yong, T S Chua* and M Kan, "Video Retrieval Using High-level features: Exploiting Query-matching and Confidence-based Weightin". *Proceedings of Conference on Image and Video Retrieval (CIVR 2006)*, ed. -, comp. - (2006): 143-152. -: Springer Verlag. (Conference on Image and Video Retrieval (CIVR 2006), 13 - 15 Jul 2006, -, Arizona, Afghanistan)

7. Setiawan*, H, M Kan* and H Li, "Ordering Phrases with Function Words". *Association of Computational Linguistics (ACL 07)* (2007). New York: Association for Computational Linguistics. (Association of Computational Linguistics, 25 - 28 Jun 2007, Prague, Czech Republic)
Publication no. 0340692

8. GOZALI*, J and M Kan, "A Rich OPAC User Interface with AJAX". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340135

9. Kan, M, "SlideSeer: A digital library of aligned document and presentation pairs". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340134

10. YAN*, S, D Lee, M Kan* and CL Giles, "Adaptive Sorted Neighborhood Methods for Efficient Record Linkage". *JCDL 2007* (2007). Vancouver: ACM Press. (Joint Conference on Digital Libraries, 18 - 23 Jun 2007, Vancouver, Canada)
Publication no. 0340133

11. Nguyen*, V and M Kan*, "Functional Faceted Web Query Analysis". *WWW 2007* (2007). Banff: ACM Press. (World Wide Web, 8 - 12 May 2007, Banff, Canada)
Publication no. 0350704

12. LIN*, Z, T S Chua, M Kan*, W S Lee, L QIU and S YE, "NUS at DUC 2007: Using Evolutionary Models of Text". *DUC 2007* (2007). Rochester: Association for Computational Linguistics. (Document Understanding Conference, 26 - 27 Apr 2007, Rochester, United States)
Publication no. 0350701

13. LIN*, Z and M Kan, "Timestamped Graphs: Evolutionary Models of Text for Multi-document Summarization". *Proceedings of TextGraphs-2* (2007). Rochester: Association for Computational Linguistics. (Graph-based Methods for Natural Language (TextGraphs-2), 26 - 27 Apr 2007, Rochester, United States)
Publication no. 0340129

14. Iskandar*, D, Y Wang, M Kan and H Li, "Syllabic Level Automatic Synchronization of Music Signals and Text Lyrics". *ACM Multimedia* (2006). New York: ACM Press. (ACM Multimedia, 23 - 27 Oct 2006, Santa Barbara, California, United States) (Rank Other (Short paper)).
Publication no. 0314882

15. Neo, Shi Yong, T S Chua* and M Kan, "Video Retrieval Using High-level features: Exploiting Query-matching and Confidence-based Weightin". *Proceedings of Conference on Image and Video Retrieval (CIVR 2006)*, ed. -, comp. - (2006): 143-152. -: Springer Verlag. (Conference on Image and Video Retrieval (CIVR 2006), 13 - 15 Jul 2006, -, Arizona, Afghanistan)

Publication no. 0328451

16. Qiu*, L, M Kan and T S Chua, "Paraphrase Recognition via Dissimilarity Significance Classification". *Coling/ACL 2006* (2006). Sydney: Association for Computational Linguistics. (Coling / ACL 2007, 17 - 23 Jul 2006, Sydney, Australia) (Rank 1).
Publication no. 0314877

17. Tan*, Yee Fan, M Kan and Dongwon Lee, "Search Engine Driven Author Disambiguation". *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (2006). New York: ACM Press. (Joint Conference on Digital Libraries, 1 - 10 Jun 2006, Chapel Hill, North Carolina, United States) ((Rank: Other. Short Paper)).
Publication no. 0306315

18. Tan*, Yee Fan, M Kan and Hang Cui, "Extending corpus-based identification of light verb constructions using a supervised learning framework". *Proceedings of the EACL 2006 Workshop on Multi-word-expressions in a multilingual context (MWEmc)* (2006). Trento: EACL. (Workshop on Multi-word-expressions in a multilingual context, 1 - 5 Apr 2006, Trento, Italy)
Publication no. 0306314

19. Nguyen Thi, Hoang Oanh and M Kan*, "Fast Webpage classification using URL features". *Proc. of Conf. on Info and Knowledge Management (CIKM '05)* (2005). New York: ACM Press. (Conf. on Info and Knowledge Management, 1 Nov 2005, Bremen, Germany) ((Poster paper. Rank: Other)).
Publication no. 0306313

20. Lu*, Wei and M Kan, "Supervised Categorization of Javascript using Program Analysis Features". *AIRS 2005* (2005). New York: Springer. (Asia Information Retrieval Symposium, 1 - 5 Oct 2005, Jeju Island, South Korea)
Publication no. 0306311

21. Cui*, H, M Kan and T S Chua, "Generic Soft Pattern Models for Definitional Question Answering". *ACM SIG on Information Retrieval (SIGIR 05)* (2005). United States. (ACM SIG on Information Retrieval (SIGIR 05), 1 Aug 2005) (Rank: 1).
Publication no. 0214339

22. Cui*, H, R Sun, K Li, M Kan and T S Chua, "Question Answering Passage Retrieval Using Depedency Relations". *ACM SIG on Information Retrieval (SIGIR 05)* (2005). United States. (ACM SIG on Information Retrieval (SIGIR 05), 1 Aug 2005) (Rank: 1).
Publication no. 0214338

23. Kan*, M and Y How, "Optimizing predictive text entry for short message service on mobile phones". *Human Computer Interfaces International (HCII 05)* (2005). United States. (Human Computer Interfaces International (HCII 05), 1 Jul 2005) (Rank: Other).
Publication no. 0214336

24. Kan*, M and C C D Poo, "Detecting and supporting known item queries in online public access catalogs". *Joint Conference on Digital Libraries* (2005). United States. (Joint Conference on Digital Libraries, 1 Jan 2005) (Rank: 2).
Publication no. 0214335

25. Lee*, C H, M Kan and S Lai, "Sylistic and Lexical Co-training for Web Block Classification". *Workshop on Web Information and Data Management (WIDM '04)* (2004). United States. (Workshop on Web Information and Data Management (WIDM '04), 1 Nov 2004, Washington, D.C., United States) (Ranking: Other).
Publication no. 0187760

26. Wang*, Y, M Kan*, T. Nwe, A. Shenoy and J Yin, "LyricAlly: Automatic Synchronization of Acoustic Musical Signals and Textual Lyrics". *ACM Multimedia 2004* (2004). New York: ACM. (ACM Multimedia 2004, 10 - 15 Oct 2004, New York, United States) (Ranking: 1).
Publication no. 0189974

27. Komarjaya*, J, C C D Poo and M Kan, "Corpus-Based Query Expansion in Online Public Access Catalogs". *European Conference on Digital Libraries (ECDL '04)* (2004). United Kingdom. (European Conference on Digital Libraries (ECDL '04), 1 Sep 2004, Bath, United Kingdom) (Ranking: Other).
Publication no. 0187755

28. Cui*, H, M Kan and T S Chua, "A Comparative Study on Sentence Retrieval for Definitional Question Answering". *Workshop on Information Retrieval for Question Answering (IR4QA)* (2004). United Kingdom. (Workshop on Information Retrieval for Question Answering (IR4QA), 1 Jul 2004, Sheffield, United Kingdom) (Rank: 2).
Publication no. 0187753

29. Qiu*, L, M Kan and T S Chua, "A Public Reference Implementation of the RAP Anaphora Resolution Algorithm". *Language Resources and Evaluation Conference 2004 (LREC 04)* (2004) (Language Resources and Evaluation Conference 2004 (LREC 04), 1 May 2004, Lisbon, Portugal) (Ranking: Other).
Publication no. 0187752

30. Cui*, H, M Kan and T S Chua, "Unsupervised Learning of Soft Patterns for Generating Definitions from Online News". *The 13th International World Wide Web Conference (WWW 2004)* (2004): 90-99. United States. (The 13th International World Wide Web Conference (WWW 2004), 1 May 2004, New York, United States) (Ranking: 1).
Publication no. 0187751

31. Cui*, H, K Li, R Sun, T S Chua and M Kan, "National University of Singapore at the TREC-13 Question Answering Main Task". *Text Retrieval Conference 13 (TREC 13)* (2004). United States. (Text Retrieval Conference 13 (TREC 13), 1 Jan 2004) (Rank: Other).
Publication no. 0214415

32. Kan, M, "Web Page Classification Without the Web Page". *13th International World Wide Web Conference (WWW2004)* (2004). United States. (13th International World Wide Web Conference (WWW2004), 1 Jan 2004, New York, United States) (Poster paper. Rank: Other).
Publication no. 0187750

33. Kushniruk*, A, M Kan, K Mckeown, J Klavans, D Jordan, M Laflamme and V Patel, "Usability Evaluation of an Experimental Text Summarization System and Three Search Engines: Implications for the Reengineering of Health Care Interfaces". *American Medical Informatics Association (AMIA 2002)* (2002). United States: American Medical Informatics Association Press. (American Medical Informatics Association (AMIA 2002), 1 Nov 2002) (Rank: 2nd Tier (Miscellaneous)).
Publication no. 0157526

34. Kan*, M, K Mckeown and J Klavans, "Domain-specific informative and indicative summarization for information retrieval". *Document Understanding Workshop (DUC 2001)* (2002). United States: ACM Press. (Document Understanding Workshop (DUC 2001), 1 Sep 2001, New Orleans, Louisiana, United States) (Ranking: Other).
Publication no. 0157531

35. Kushniruk*, A, M Kan, K Mckeown, J Klavans and V Patel, "Evaluating the Content and Usability of an Experimental Text Summarization System and Three Web-Based Search Engines". *Human Factors and Ergonomics 46th Annual Meeting (HFES 2002)* (2002). United States. (Human Factors and Ergonomics 46th Annual Meeting (HFES 2002), 1 Sep 2002) (Ranking: Other).
Publication no. 0157527

36. Kan*, M and K Mckeown, "Corpus-trained text generation for summarization". *International Workshop on Natural Language Generation (INLG 2002)* (2002): 1-8. United States: Association for Computational Linguistics. (International Workshop on Natural Language Generation (INLG 2002), 1 Jul 2002, Harriman, New York, United States) (Ranking: Other).
Publication no. 0157529

37. Kan*, M and J Klavans, "Using Librarian Techniques in Automatic Text Summarization for Information Retrieval". *Joint Conference on Digital Libraries (JCDL 2002)* (2002): 39-45. United States: ACM Press. (Joint Conference on Digital Libraries (JCDL 2002), 1 Jun 2002, Portland, Oregon, United States) (Ranking: 2nd Tier (Applications and Media)).
Publication no. 0157528

38. Kan*, M, J Klavans and K Mckeown, "Using the Annotated Bibliography as a Resource for Indicative Summarization". *Language Resources and Evaluation Conference (LREC 2002)* (2002): 1746-1752. Spain: ELRA. (Language Resources and Evaluation Conference (LREC 2002), 1 May 2002, Las Palmas, Canary Islands, Spain) (Ranking: Other).
Publication no. 0157530

39. Kan*, M, M Mckeown and J Klavans, "Applying Natural Language Generation to Indicative Summarization". *European Workshop on Natural Language Generation* (2001): 92-100. France. (European Workshop on Natural Language Generation, 1 Jul 2001) (Ranking: Other).
Publication no. 0166953

40. Hatzivassiloglou*, V, J Klavans, M Holcombe, R Barzilay, M Kan and K Mckeown, "Simfinder: A Flexible Clustering Tool for Summarization". *Proceedings of the Workshop on Summarization in NAACL `01* (2001). Pittsburgh, USA. (Proceedings of the Workshop on Summarization in NAACL `01, 1 Jun 2001) (Ranking: Other).
Publication no. 0166964

41. Mckeown*, K, R Barzilay, D Evans, V Hatzivassiloglou, M Kan, B Schiffman and S Teufel, "Columbia Multi-Document Summarization: Approach and Evaluation". *Document Understanding Workshop (2001)* (2001). United States: National Institutes of Standards and Technology (NIST). (Document Understanding Workshop (2001), 1 Jan 2001, New Orleans, United States) (Ranking: Other).
Publication no. 0166949

42. Kan*, M, J Klavans and K Mckeown, "Linear Segmentation and Segment Relevence". *Proceedings of 6th International Workshop of Very Large Corpora (WVLC-6)* (1998): 197-205. Quebec, Canada. (Proceedings of 6th International Workshop of Very Large Corpora (WVLC-6), 1 Aug 1998) (Ranking: Other (WVLC became EMNLP, currently rank 2)).
Publication no. 0166971

43. Klavans*, J and M Kan, "Role of Verbs in Document Analysis". *Proceedings of COLING/ACL 98* (1998): 680-686. Quebec, Canada. (Proceedings of COLING/ACL 98, 1 Aug 1998) (Ranking: 1/2 (COLING is rank 2, ACL is rank 1; the conference was a joint conference)).
Publication no. 0166965

44. Klavans, J, K Mckeown, M Kan and SUSAN Lee, "Resources for the Evaluation of Summarization Techniques". *Proceedings of the 1st International Conference on Language Resources and Evaluation* (1998). Grenada, Spain. (Proceedings of the 1st International Conference on Language Resources and Evaluation, 1 May 1998) (Ranking: Other).
Publication no. 0166968

45. Fung, P, M Kan and Y Horita, "Extracting Japanese Domain and Technical Terms is Relatively Easy". *Second International Conference in New Methods for Language Processing, (NEMLP)* (1996): 148-159. Bilkent, Turkey. (Second International Conference in New Methods for Language Processing, (NEMLP), 1 Sep 1996) (Ranking: Other).
Publication no. 0166975

POSTER PRESENTATION

PUBLISHED

46.   Wang, F and M Kan*, "NPIC: Hierarchical synthetic image classification using image search and generic features". *Conference on Image and Video Retrieva* (2006). New York: ACM Press. (Conference on Image and Video Retrieval, 10 - 15 Jul 2006, ASU, Tempe, United States)
Publication no. 0310626

47.   SUN, R.X., H CUI, M Kan and T S Chua*, "Dependency relation matching for answer selection". *Conference on Research and Development in information Retrieval, SIGIR '05* (2005): 651-652. New York: ACM. (Conference on Research and Development in information Retrieval, SIGIR '05, 15 - 19 Aug 2005, Hotel, Salvador, Brazil)
Publication no. 0327986

EDITORIAL WORK ON BOOK

PUBLISHED

1.   Ng*, H T, M-K Leong, M Kan and D Ji, eds., *Information Retrieval Technology. Proceedings of the Third Asia Information Retrieval Symposium, AIRS 2006*, Berlin Heidelberg: Springer, 2006. 683 pp.
Publication no. 0342296

CHAPTER IN BOOK

PUBLISHED

1.   Kan, M, "Using multi-document summarisation to assist in semi-structured literature retrieval: A case study in consumer healthcare". In *Design and Usability of Digital Libraries : Case Studies in the Asia Pacific*, ed. Theng, Y.L. and Foo, S.. IDEA Group, 2005.
Publication no. 0189194

THESIS

UNPUBLISHED BUT FINALIZED

1.   Kan, M, *Automatic text summarization as applied to information retrieval: Using indicative and informative summaries*. Ph.D. in Computer Science thesis, Columbia University, New York, New York, United States. 2003. 336 pp. ((available from http://www.cs.columbia.edu/~min/papers/thesis.pdf)).
Publication no. 0157163

PUBLISHED REPORT

PUBLISHED

1.   Gozali*, J and M Kan*, *Rich and Dynamic Library Catalogs: A Case Study of Online Search Interfaces*. Technical Report, no. A8/07. Singapore: National University of Singapore, 2007. 10 pp.

Publication no. 0350703

2.   Kan*, M and H Nguyen thi, *Fast webpage classification using URL features*. Technical Report, no. TRC 08/05. SINGAPORE, 2005.
Publication no. 0214408

3.   Tan*, Y, M Kan and H Cui, *Extending corpus-based identification of light verb constructions using a supervised learning framework*. Technical Report, no. TRB 08/05. SINGAPORE, 2005.
Publication no. 0214406

4.   Yang*, H, H Cui, M Kan, M Maslennikov, L Qiu and T S Chua, *QUALIFIER in TREC 12 QA Main Task*. TREC Notebook paper. Washington D.C.: National Institute of Science and Technology, 2003. 10 pp.
Publication no. 0174437

5.   Kan, M, *Metadata extraction and text categorization using Universal Resource Locator expansions*. Technical Report, no. TR 10/03. Singapore: National University of Singapore Department of Computer Science, 2003. 8 pp.
Publication no. 0174340

6.   Kan, M, *Combining visual layout and lexical cohesion features for text segmentation*. Technical Report, no. CUCS-002-01. New York, NY: Columbia University, 2001.
Publication no. 0166981

7.   Kan*, M, J Klavans and K Mckeown, *Synthesizing composite topic structure trees for multiple domain specific documents*. Technical Report, no. CUCS-003-01. New York, NY: Columbia University, 2001.
Publication no. 0166979

8.   Braschler, M, M Kan, P Schauble and J Klavans, *The Eurospider Retreival System and the TREC-8 Cross-Language Task*. Technical Report, no. NIST Special Publication 500-246. Gaithersburg, MD: National Institutes of Standards and Technology, 1999.
Publication no. 0166990

9.   Kan*, M and K Mckeown, *Information Extraction and Summarization: Domain Independence through Focus Types*. Technical Report, no. CUCS-030-99. New York, NY: Columbia University, 1999.
Publication no. 0166985